"I hereby declare that I have read this dissertation and in my opinion this thesis has satisfied the scope and quality for the award of the degree of Master of Science (Biotechnology)."

Signature: ……..…….
Name of supervisor: Dr SHAHIR SHAMSIR
Date: ……...

STRUCTURAL ANALYSIS AND MODELING COMPARISON OF PRIMATES' DYSTROPHIN

FAEZE SADAT MOHAJER

A dissertation submitted in partial fulfilment of the

requirements for the award of the degree of

Master of Science (Biotechnology

Faculty of Bioscience and Bioengineering

Universiti Technologi Malaysia

January 2013

I declare that this thesis entitled "**Structural analysis and modelling comparison of primates' Dystrophin"** is the result of my own research except as cited in the references. The thesis has not been accepted for any degree and is not concurrently submitted in candidature of any other degree.

Signature: .....

Name: Faeze Sadat Mohajer.....

Date: Dec .2012.....

To my beloved parents, sisters and brother

ACKNOWLEDGMENT

I would like to begin by sincerely thanking my supervisor, Associated Prof Dr. Shahir Shamsir for his kindness and constant support, guidance and mentorship over the course of this thesis. He gave me the freedom to define my thesis statements. I am very grateful to having the opportunity for work on this thesis.

I am heartily thankful to my parents, for their unconditional supports and encouraging me to continue my education.

ABSTRACT

Dystrophin is a rod-shaped cytoplasmic protein which is an essential piece of a protein complex that binds the cytoskeleton of a muscle fiber to the nearby extracellular matrix within the cell membrane. Lack of this protein in the muscle leads muscular dystrophy. Currently, Dystrophin structure has just been detected in human being. In fact, it has not been recognized in other types of primates. However, bioinformatics' software has presented a big opportunity to predict its structure in other species on based on its structure in human being. The data for this project will be sourced by both bioinformatics databases and experimental information. The used databases were NCBI, PDB (protein data bank), UniProt, and Gene bank. The data that is gathered from these databases were analyzed using multiple sequence analysis software such as BLAST, Jalview to find conserved regions in the protein among the species of interest (different kinds of primates). In addition, the tertiary structure of the protein will be predicted by Swiss-model and will be observed and compared by 3-d structure viewing software such as: VMD and DeepView. In conclusion, the analysis of this study illustrates that Dystrophin has been changed slightly through evolution. However, the small changes that have been occurred have not affected the 3D structure and function of the protein.

ABSTRAK

Dystrophin adalah protein berbentuk rod cytoplasmic yang merupakan sekeping penting kompleks protein yang mengikat cytoskeleton gentian otot untuk matriks extracellular dalam membran sel berdekatan. Kekurangan protein ini dalam otot membawa distrofi otot. Pada masa ini, struktur Dystrophin baru sahaja telah dikesan dalam manusia. Malah, ia telah tidak diiktiraf dalam lain-lain jenis primat. Walau bagaimanapun, perisian bioinformatik 'telah dibentangkan peluang yang besar untuk meramalkan struktur dalam spesies lain berdasarkan struktur dalam manusia. Data untuk projek ini akan diperolehi oleh kedua-dua pangkalan data bioinformatik dan maklumat eksperimen. Pangkalan data yang digunakan adalah NCBI, PDB (protein data bank), UniProt, dan Gene bank. Data yang dikumpul daripada pangkalan data ini dianalisis menggunakan pelbagai urutan perisian analisis seperti letupan, Jalview untuk mencari kawasan terpelihara dalam protein kalangan spesis faedah (jenis primat). Di samping itu, struktur pengajian tinggi protein akan diramalkan oleh Swiss model dan akan diperhatikan dan berbanding oleh 3D tontonan perisian struktur seperti: VMD dan DeepView. Kesimpulannya, analisis kajian ini menunjukkan bahawa Dystrophin telah berubah sedikit melalui evolusi. Walau bagaimanapun, perubahan kecil yang telah berlaku tidak terjejas 3D struktur dan fungsi protein.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

DMD	Duchenne Muscular Dystrophy
BMD	Becker Muscular Dystrophy
ECM	Extra Cellular Matrix
3-D structure	Three Dimensional Structure
NCBI	National Center for Biology Information
PDB	Protein Data Bank
Mya	Million Years ago
OWM	Old World Monkey
NWM	New World Monkey
PrP	Prion Protein
GHR	Growth Hormone Receptor
UniProt	Universal Protein Resource
BLAST	Basic Local Alignment Search Tool
MSA	Multiple Sequence Alignment
VMD	Visual Molecular Dynamic
MD	Molecular Dynamic
RMSD	Rootmean- Square Deviation
DAG	Dystrophin-Associated Glycoprotein
NMR spectroscopy	Nuclear magnetic resonance spectroscopy
QMEAN4	Qualitative Model Energy ANalysis

LIST OF SYMBOLS

aa	Amino Acid
Asp	Aspartic acid
Cys	Cysteine
Ser	Serine
Ala	Alanine
Gly	Glycine
Glu	Glutamic acid
Gln	Glutamine
Asn	Asparagine
Lys	Lysine
Arg	Arginine
Met	Methionine
Ala	Alanine

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Dystrophin is a rod-shaped protein that is found in both skeletal and cardiac muscles. It is mainly a structural protein with different variety that is suited to different types of organs in the human body. Dystrophin is mostly located in muscles that are used for movement (skeletal muscles) and the muscles of the heart (cardiac muscles). There are small amount of dystrophin in nerve cells of the brain (Chelly, et al., 1990). Dystrophin also works as a part of a group of proteins that interact together in a protein complex called the costamere or the dystrophin-associated protein complex. This protein complex is responsible for the strengthening of muscle fibres and protects them from injury during muscles contraction and relaxation. The dystrophin protein is encoded in the DMD gene and is sometimes called the DMD protein due to its association with the Duchenne Muscular Syndrome (DMD) disease. In muscle, the DMD complex acts as an anchor that connects each one of the muscle cell's structural framework (cytoskeleton) with the lattice of proteins and other molecules outside the cell (extracellular matrix). In fact, it connects actin of cytoskeleton by its N-terminal (action binding site) to ECM (Extra Cellular Matrix). (Ozawa, et al., 1995). In addition, it may also have a

responsibility in cell signalling with its interaction with proteins that send and receive chemical signals. Currently, there is only limited knowledge on the role of this protein in nerve cells. Research has proposed that dystrophin in nerve cells are necessary for normal structure and function of synapses, which are important for specializing connections between nerve cells in the place that cell-to-cell communication happens.(Ahn & Kunkel, 1993)

1.1.1 Duchenne Muscular Dystrophy (DMD)

Normal skeletal muscle is composed of muscle fibres that are equally spaced, sharp, and has a moderately uniform size. They are symmetric, and also are multinucleated with nuclei's which are placed at the periphery of the skeletal fibre. In DMD, patient that lacks dystrophin shows the muscle cells of patients who are suffered from Duchenne Muscular Dystrophy (DMD) are abnormally defenceless. These suggests that dystrophin plays a significant role in muscle rigidity, making muscle cells possessing dystrophin being far stiffer than cells which lack dystrophin (Pasternak, Wong, & Elson, 1995). Fatal DMD muscle is quite normal except for the occasions that fibres are eosinophilic hypercontracted (Blake, Weir, Newey, & Davies, 2002). DMD gene locus is on X chromosome. Thus, Duchenne Muscular Dystrophy (DMD) is an X-linked recessive mutation where approximately 1 in 3,500 boys are suffered from this disease (Blake, et al., 2002). Patients suffer from muscle waste, usually lose their ability for walking and are paralyzed before the age of 12 and die in their late teens or early twenties typically because of respiratory failure. A milder type for this disease is called the Becker muscular dystrophy (BMD), which symptoms are detected later and the patient survival time is longer. The most extensive mutations can cause the complete dystrophin deficiency, while a truncated protein presents in a very limited level are found in BMD patients. In addition, mutations in the genes can affect the encoding of

many constituents of the dystrophin associated protein complex which can cause other forms of this disease like the limb-girdle muscular dystrophies and congenital muscular dystrophy (Blake, et al., 2002).

1.2 Problem Statement

Currently, research has only determined the dystrophin structure for humans. None of the structure from primates has been determined either experimentally or computationally using protein structure modelling. A comparison between human and primate dystrophin will elucidate its evolutionary path, especially in its most significant domain (the first 246 amino acids of its N-terminal (actin-binding site) of the protein) (Pasternak, et al., 1995). The aim of this study is to compare the dystrophin structure between human and primates by first modelling the structure and analysing the differences in the amino acid composition, 3D structure and evolutionary changes.

1.3 Objectives

The objectives of this study are:

1. To model the structure of dystrophin from primates for comparison with human dystrophin.
2. To analyse the differences in the amino acid sequences and composition between different primates and human.

3. To analyse the 3D structure and investigation the implication of the differences in amino acid sequence of dystrophin between human and primate.

1.4 Scope of the Study

The data for this project will be sourced by bioinformatics databases that stores experimental information such as the NCBI (http://www.ncbi.nlm.nih.gov), the Protein Data Bank (www.rcsb.org), the UniProt (http://www.uniprot.org) and the Genbank (http://www.ncbi.nlm.nih.gov/genbank/).

REFRENCES

Ahn, A. H., & Kunkel, L. M. (1993). The structural and functional diversity of dystrophin. Nature genetics, 3(4), 283-291.

Baumbach, L., Chamberlain, J., Ward, P., Farwell, N., & Caskey, C. (1989). Molecular and clinical correlations of deletions leading to Duchenne and Becker muscular dystrophies. Neurology, 39(4), 465-465.

Blake, D. J., Weir, A., Newey, S. E., & Davies, K. E. (2002). Function and genetics of dystrophin and dystrophin-related proteins in muscle. Physiological reviews, 82(2), 291-329.

Chelly, J., Hamard, G., Koulakoff, A., Kaplan, J. C., Kahn, A., & Berwald-Netter, Y. (1990). Dystrophin gene transcribed from different promoters in neuronal and glial cells, 344, 64 - 65

Cheng, Z., Gengbin, S., Zhuolin, L., Xiulin, L., & Jianhua, C. Correlation between the hydrophobic structure of dystrophin and DMD/BMD, 12(6):341-344]

Diogo, R., Richmond, B. G., & Wood, B. (2012). Evolution and homologies of primate and modern human hand and forearm muscles, with notes on thumb movements and tool use. Journal of human evolution, 32(7), 64-78

Gilad, Y., Oshlack, A., Smyth, G. K., Speed, T. P., & White, K. P. (2006). Expression profiling in primates reveals a rapid evolution of human transcription factors. Nature, 440(7081), 242-245.

Hasegawa, M., Kishino, H., & Yano, T. (1985). Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of molecular evolution, 22(2), 160-174.

Hu, X., Burghes, A., Ray, P. N., Thompson, M. W., Murphy, E., & Worton, R. G. (1988). Partial gene duplication in Duchenne and Becker muscular dystrophies. Journal of medical genetics, 25(6), 369-376.

Labuda, D., Zietkiewicz, E., & Yotova, V. (2000). Archaic lineages in the history of modern humans. Genetics, 156(2), 799-808.

Lenk, U., Hanke, R., Kräft, U., Grade, K., Grunewald, I., & Speer, A. (1993). Non-isotopic analysis of single strand conformation polymorphism (SSCP) in the exon 13 region of the human dystrophin gene. Journal of medical genetics, 30(11), 951-954.

Luther, P. K., Squire, J. M., & Forey, P. L. (1996). Evolution of myosin filament arrangements in vertebrate skeletal muscle. Journal of morphology, 229(3), 325-335.

Muthu, M., Richardson, K. A., & Sutherland-Smith, A. J. (2012). The Crystal Structures of Dystrophin and Utrophin Spectrin Repeats: Implications for Domain Boundaries. PloS one, 7(7), 23-56.

Norwood, F., Sutherland-Smith, A. J., Keep, N. H., & Kendrick-Jones, J. (2000). The structure of the N-terminal actin-binding domain of human dystrophin and how mutations in this domain may cause Duchenne or Becker muscular dystrophy. Structure (London, England: 1993), 8(5), 481-564.

Oshima, J., Magner, D. B., Lee, J. A., Breman, A. M., Schmitt, E. S., White, L. D., et al. (2009). Regional genomic instability predisposes to complex dystrophin gene rearrangements. Human genetics, 126(3), 411-423.

Ozawa, E., Yoshida, M., Suzuki, A., Mizuno, Y., Hagiwara, Y., & Noguchi, S. (1995). Dystrophin-associated proteins in muscular dystrophy. Human molecular genetics, 4, 1711-2345.

Pantel, J., Machinis, K., Sobrier, M. L., Duquesnoy, P., Goossens, M., & Amselem, S. (2000). Species-specific Alternative Splice Mimicry at the Growth Hormone Receptor Locus Revealed by the Lineage of Retroelements during Primate Evolution a novel mechanism acounting for protein diversity

between and within species. Journal of Biological Chemistry, 275(25), 18664-18669.

Pasternak, C., Wong, S., & Elson, E. L. (1995). Mechanical function of dystrophin in muscle cells. The Journal of cell biology, 128(3), 355-361.

Schätzl, H. M., Da Costa, M., Taylor, L., Cohen, F. E., & Prusiner, S. B. (1995). Prion protein gene variation among primates. Journal of molecular biology, 245(4), 362-564.

Sun, X., Zhulin,I. & Wartell, R. M. (2002). Predicted structure and phyletic distribution of the RNA-binding protein Hfq. Nucleic acids research, 30(17), 3662-3671.

Takahata, N., & Satta, Y. (1997). Evolution of the primate lineage leading to modern humans: phylogenetic and demographic inferences from DNA sequences. Proceedings of the National Academy of Sciences, 94(9), 4811-4815-5985.

Takeshima, Y., Yagi, M., Okizuka, Y., Awano, H., Zhang, Z., Yamauchi, Y., et al. (2010). Mutation spectrum of the dystrophin gene in 442 Duchenne/Becker muscular dystrophy cases from one Japanese referral center. Journal of human genetics, 55(6), 379-388.

Yan, J., Wen, W., Xu, W., Long, J., Adams, M. E., Froehner, S. C., et al. (2005). Structure of the split PH domain and distinct lipid-binding properties of the

PH–PDZ supramodule of α-syntrophin. The EMBO journal, 24(23), 3985-3995.

Zhang, J., Webb, D. M., & Podlaha, O. (2002). Accelerated protein evolution and origins of human-specific features: Foxp2 as an example. Genetics, 162(4), 1825-1835.